DOCUMENT RESUME

ED 225 349 FL 013 369

AUTHOR Von Elek, Tibor
TITLE Test of Swedish as a Second Language: An Experiment in Self-Assessment. Work Papers from the Language Teaching Research Center, No. 31.
INSTITUTION Gothenburg Univ. (Sweden). Language Teaching Research Center.
PUB DATE May 82
NOTE 33p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adults; Autoinstructional Aids; Immigrants; Language Proficiency; Language Skills; *Language Tests; *Second Language Learning; *Self Evaluation (Individuals); *Swedish; Testing
IDENTIFIERS Sweden

ABSTRACT

A self-assessment test of Swedish as a second language is described. The test was designed as a self-diagnostic instrument for adult migrants to Sweden studying Swedish at labor training centers. The 1,500-item test provides students with information about their Swedish proficiency level as well as the proficiency needed for their professional training. The test items are arranged under six skill areas: vocabulary, grammar, listening comprehension, reading comprehension, oral proficiency, and written proficiency. In each skill area, students rate their own understanding of each item. Thus, students can diagnose their specific weaknesses in relation to their proficiency requirements and work to correct them. The first version of the test was administered to 300 adult migrants enrolled in Swedish courses. An oral followup procedure was conducted by the teachers to verify whether the subjects actually knew the items they claimed to know in the self-assessment procedure. The results showed that the students' self-judgments were highly reliable. The students reported that they enjoyed the test and viewed it as a non-threatening teaching aid rather than a test. Sample items from the test are appended. (RW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

SPRÅKPEDAGOGISKT CENTRUM
GÖTEBORGS UNIVERSITET
BOX 1010, 431 26 MÖLNDAL
TEL. 031-87 80 00

ED225349

TIBOR VON ELEK

TEST OF SWEDISH AS A SECOND LANGUAGE: AN EXPERIMENT IN SELF-ASSESSMENT

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Language Teaching Research Center

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

FL 013 369

WORK PAPERS FROM THE LANGUAGE TEACHING RESEARCH CENTER
No. 31 May 1982
ISSN 0349-2486

TABLE OF CONTENTS

1 Background to the project 1

2 Design, format and use of the test 4

3 Content and levels of difficulty 10

4 The try-out run 12

5 Results and experience 15

6 Some problems and future plans 18

References 20

Appendix Sample Items from the second version of the self-assessment test 21

TEST OF SWEDISH AS A SECOND LANGUAGE: AN EXPERIMENT IN SELF-ASSESSMENT

The purpose of this paper is to present a preliminary version of a self-assessment test of Swedish as a second language.[1] It is being developed at the Language Teaching Research Center, University of Göteborg, within a project sponsored by the National Swedish Board of Education. The test is meant to serve as a self-diagnostic instrument for adult migrants studying Swedish as a second language. It is hoped that, if experimentation proves successful, the test will provide a model easily adaptable to other second or foreign languages.

1 Background to the project

During the 1970's the teaching of Swedish to migrants developed into a school subject which in scope is second only to English among modern languages taught in Sweden. The rapid development of Swedish as a second language has led to acute needs for teaching aids, teacher training programs, and, of course, various test instruments.

The urgent need for placement and achievement tests has mainly been met by tests constructed by teachers. To a certain extent these tests are also diagnostic instruments in that they give the teacher some information about linguistic areas within which the student needs to improve his ability. The majority of these teacher-made tests could probably be classified as "pre-scientific" or "traditional" types of tests in Spolsky's terms (1976).

A proficiency test which presumably would qualify as a representative of Spolsky's "psychometric-structuralist" or "modern" period (Ibid.) was developed at the Language Teaching Research Center in 1974-75. The test was to a certain degree modelled on the TOEFL, and it was primarily developed for the labour market training schools where large

1) The present paper is an abbreviated and updated version of a Swedish report (von Elek, 1981).

numbers of adult migrants attend courses in Swedish. Scores from this test are "translated" into proficiency levels on a rating scale defining, in quite general terms, the testee's global proficiency in Swedish. These test results often underlie decisions about the student's admission to professional training or to additional courses in Swedish. The test has been described in detail in a report which also presents the findings of a quite extensive investigation into the migrants' proficiency in Swedish as measured by means of the same test (von Elek, 1977).

A common complaint on the part of adult migrants receiving their training in Swedish at labour market centers and awaiting subsequent professional training or re-training, has been that they lack information about their actual Swedish proficiency and also information about the level they are supposed to achieve in order to qualify for and be able to cope with the job or professional training they are aiming at. What they have in mind is not a vague and general description of these proficiency levels but quite concrete and explicit information on their own proficiency profile in relation to the language proficiency they ought to achieve - or at least approach - in the course of their studies. Such information, they believe, would make their course work more goal-oriented and would consequently motivate and stimulate them in their efforts to make up for obvious shortcomings.

These expectations imply that in the teaching of a second language to migrants, especially adults, conventional achievement tests and end-of-course tests are of limited value. Instead, emphasis should be put on the development of instruments that will enable the students themselves to diagnose their weak areas and organise their studies with a view to satisfying individual needs.

The guiding principle in our assignment within the project has, therefore, been to try to provide migrants with the type of instrument that would correspond to their expectations

mentioned above. It was also decided that this instrument should be constructed according to self-assessment principles. There are several reasons for this.

Experiments with self-assessment instruments, some of them conducted at the Language Teaching Research Center by Oskarsson (1980) and von Elek (1978), have shown that adults are fully capable of making reliable judgements about their own mastery of a foreign language, provided they have access to adequate instruments. Therefore, there is good reason to give adult students an opportunity to assume greater responsibiliby in the evaluation process and thereby to contribute to their autonomy. Apart from these psychological and motivational aspects the engagement of students in self-assessment has also some practical implications. It allows the development of tests with more substantial, more varied and, consequently, more valid content. It also reduces the teacher's burden of marking, an activity that does not contribute to anybody's learning and can therefore be regarded as a waste of time.

We also wish to see our assignment as an opportunity to take a step from the "modern" period towards Spolsky's third, or "post-modern" period, i e the "sociolinguistic-psycholinguistic" one. Thus instead of overemphasising test reliability we are trying to give priority to high validity. This we hope to achieve partly by providing subtests in all skills, including the productive ones, and partly by playing down the role of descrete point items in favour of integrated tests. This also means transferring the emphasis from linguistic ability to communicative competence.

Finally, by introducing our test, which we should like to classify as a representative of "the gentle art of diagnostic testing", we hope to show that testing does not necessarily have to be a dramatic and formal event in a tense atmosphere of teacher authority and student anxiety. It will probably not be possible to eliminate formal tests altogether in the

future, or at least as long as we need test results underlying end-of-term statements, grades or ranking. We should, however, like to see them used alternatively with informal, student-centered tests that can be taken as relaxed classroom activities or homework, and in which the teacher's role is that of a helper and partner assisting the learner to make the best possible use of the information provided by his self-assessment.

In summing up it can be said that the aim of our work is to develop an instrument which will enable the target group, i e adult migrants studying Swedish in order to become full-fledged citizens in their new country,

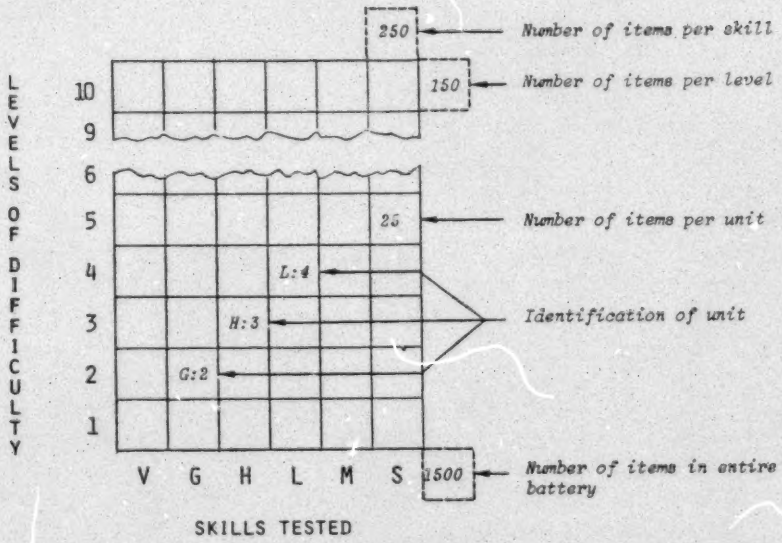
- to assume greater responsibility in the evaluation of their proficiency and progress,
- to diagnose their weak areas and obtain a realistic view of their general proficiency as well as their skills profile,
- to see their actual proficiency in relation to the level they wish to achieve in order to qualify for a certain job or training program, and consequently
- to become more motivated and goal-oriented in their further studies.

2 Design, format and use of the test

Our test consists of a large number of items (1,500) arranged into content areas (hereafter: skills) on the one hand and into difficulty levels on the other. The grid in Figure 1 illustrates the structure of the entire battery. It falls into six skills (vocabulary, grammar, listening and reading comprehension, oral and written production) and within each of these into ten difficulty levels. Thus the battery encompasses a total of 60 subtests or units, each of which comprises 25 items.

Figure 1

The structure of the self-assessment test



Letters (initials of corresponding Swedish terms) stand for the following content areas:

V = Vocabulary H = Listening compr. M = Oral prof.
G = Grammar L = Reading compr. S = Written prof.

A test of these dimensions is only practicable if it can be self-administered and self-scored. To make this possible the items are presented in sections headed by questions like: Do you understand the underlined word? Can you give the opposite of the underlined word? Do you understand this sentence/passage? etc. The student is asked to give his answer by checking one of the following alternatives for each item: 1) Yes, absolutely, 2) I think so, 3) No.

Normally the student starts off at the lowest level or, if he is not a beginner, at some other suitable level. He is

supposed to do all six units at the same difficulty level before he proceeds to the next one. If he has checked alternative 1 for a minimum of 80% of the items (i.e. 20 out of 25) in a unit, he can consider himself to have passed the level in question. If his "yes-answers" are below 50% in a unit there is no point in going on to the next level within that particular skill.

In Figures 2 and 3 we present the results of an imaginary student's self-assessment at levels 1-5. Figure 2 shows the percentage of items answered with "yes" within each unit. In Figure 3 the units with 80% or more of such answers have been shaded so as to make the student's skills profile more apparent. In this particular case the student seems to be comparatively poor at grammar and writing but these weaknesses are compensated for by a high level in oral skills. His average proficiency corresponds to level 3, which is two levels short of the student's imaginary goal-line at level 5. The most important information gained from the outcome of this assessment is that our student ought to concentrate on grammar and writing, areas in which the distance between the actual and desired levels is the greatest.

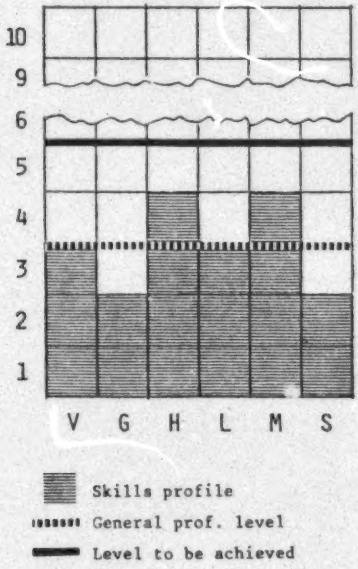
Figure 2

Items answered with 'Yes' (percentages) in an imaginary student's self-assessment

	V	G	H	L	M	S
10						
9						
6						
5	60	50	70	60	70	50
4	70	60	80	70	80	60
3	80	70	90	80	90	70
2	90	80	100	90	100	80
1	100	90	100	100	100	90

Figure 3

The outcome of the self-assessment presented in Figure 2



As is apparent the diagnostic information is given in the form of a skills profile, i.e. in terms of the relative status of the student's various skills. Therefore our test probably does not conform to conventional diagnostic tests, i.e. instruments by means of which teachers can find out what sorts of errors their students make, particularly in the fields of pronunciation, grammar and vocabulary. We believe that such diagnostic instruments can and should be constructed by the teachers themselves in accordance with the relevant syllabus, their knowledge of the class, and other factors. A diagnostic test developed by professionals for the entire target population should instead aim at revealing weaknesses in a wider sense, i.e. in terms of skills. For this reason it would probably be more adequate to classify our test as a diagnostically oriented proficiency test.

We think that especially in the teaching of migrants this type of diagnosis is essential, since teachers of Swedish as a second language seem to agree that different migrant groups demonstrate rather deviating profiles depending on such background variables as educational and cultural background, time spent in Sweden, sex (for certain groups) etc. It is desirable that both the student and the teacher should be made aware of the student's specific profile at an early stage of the course, so that adequate measures can be taken for an individualised remedial study program. It should be pointed out that not only the student's proficiency but also the "goal-line" can be expressed in the form of a specific profile, especially for careers where certain language skills are of particular importance and therefore required more than others.

Our use of a graded self-assessment test is not unlike the procedure employed in checking our eye-sight by means of a chart containing letters or texts of decreasing size. The person with a visual defect starts off by reading the largest letters or texts and proceeds to the smaller ones. The level where he gets stuck enables the optician - and to a certain extent the patient himself - to make inferences about his eye-sight. Once the patient's eye-sight ("student's proficiency") has been diagnosed the optician ("teacher") prescribes adequate glasses ("remedial measures"). In doing so he takes into consideration the patient's visual deficiency for both eyes ("profile") on the one hand, and the purpose for which he needs the glasses, e.g. reading, driving a car, etc. ("goal line") on the other.

The question has sometimes been raised why we offer the students three alternative answers to choose from if there is only one that counts. Our motive for giving three options is psychological rather than technical. We believe that students would be reluctant to check the No-alternative in cases when they are almost certain they know the correct answer, and would find it more satisfactory to be able to

answer "I think so" even though this does not affect the outcome of their self-assessment. Furthermore, eliminating this option would invite some students to check the "Yes"-box even when they are not absolutely sure of the correct answer, and this would lower the reliability of the test.

Another question we are frequently asked is whether we can rely on the students' frankness in checking the proper alternative and/or their capability of doing so. We have to admit, partly on the basis of experience gained so far, that this self-diagnostic procedure does not suit all individuals equally and that during the try-out period several students were found to have made misleading self-assessments. Still, we firmly believe that most students will give honest answers as long as they know they will neither be rewarded nor penalized for the outcome of their "confession", and as long as they know that they will deceive no one but themselves by intentionally over-estimating or underrating their capacity, and that in the long run they will benefit from a remedial program designed on the bases of realistic self-assessment results. Parallelly, nobody would wish to "deceive" the optician even if this were possible. We are convinced that in the majority of cases unrealistic self-assessments are due to insufficient information about the real purpose of the test, and perhaps to the student's deep-rooted ambition to achieve, on any test, as high scores as he possibly can.

To the question whether we can expect students to be capable of making reliable judgements about their mastery of the test-content we reply with another question: Who should know better that the student himself whether he understands a word or sentence, whether he would be able express himself adequately in a given situation, etc? We wish to point to the fact that by self-assessment we do not mean sophisticated descriptions of the student's own proficiency or profile. We do not even ask the student to choose among several descriptions of proficiency levels the one that would best

correspond to his own level, a task that may be quite demanding for people with modest educational or linguistic backgrounds. The only thing the student is supposed to do is to state honestly the degree of his knowledge of specific items according to the above mentioned alternative answers. Thanks to thorough developmental work, the interpretation of his accumulated Yes-answers is built into the instrument and will be a mere routine procedure both for the student himself and for the teacher. It is true that some items are easier to make judgements about than others and in some instances, owing to misunderstanding, the student may unintentionally give false information, but we believe that these are exceptions rather than the rule.

Another question that may be justified is this one: How are we to know where to draw the proper "goal line" for different purposes. The answer is, at this stage we do not have an exact knowledge of what our difficulty levels mean in explicit terms, which also explains why we have limited ourselves to merely assigning numbers to our levels. However, once the final version of the test has been developed and used for a certain time, it will be quite possible through experience and by means of correlation studies to determine the "goal lines" or "goal profiles", at least for the most frequent and important purposes. For the time being our "goal lines" are hypothetical and we merely use them as a demonstration of how the test is meant to be used.

3 Content and levels of difficulty

So far very little has been said about test content. One of the advantages of this type of test is that as far as choice of subject matter is concerned it is free from the limitations of conventional, especially multiple choice tests. This means that practically any linguistic material can be included, which of course is of great importance for test validity. Our guiding principle in constructing items

has been to use as much authentic material as possible. We have also been anxious to distribute the language samples to subtests in which they would fit in naturally. Thus the items that ended up in the listening comprehension units consist of samples of everyday spoken language: at lower levels simple questions, statements and instructions, more than once containing commonly used slang expressions, and at the higher levels excerpts from broadcast material, messages through loudspeakers, etc. Similarly, the reading comprehension units contain the kind of language samples that are likely to reach the migrants in written form: signs, messages, headlines, excerpts from brochures and the press, etc. The oral and written parts were constructed according to similar criteria. In the vocabulary and grammar parts we could not follow the same principles. We selected words and grammar structures that experienced teachers regarded as relevant and important. We believe that, although the test content was not compiled on the basis of a systematic needs analysis, the substantial material comprised in 1,500 items constitutes a representative random sample of the Swedish language.

Another important task besides selecting subject matter has been the distribution of items to various difficulty levels. It is not enough to grade items so that the units with higher numbers should be increasingly difficult. It is just as important to achieve a balance horizontally, i.e. to make the various subtests of the same level equally difficult. In the first version of our test this had to be done by means of the subjective judgements of a number of highly qualified and experienced teachers. In the second version, however, we could use the statistical data gained from the first try-out.

4 The try-out run

The first experimental version of the self-assessment test battery was tried out during the academic year of 1981/82. It was given to roughly 300 adult migrants enrolled in various courses of Swedish. They represented a wide range of proficiency levels from beginners to near-native speakers.

The teachers who engaged themselves in the project had a very important part to play in the try-out. They were to select subjects from their classes, distribute and correct test material, assist students with instructions, complete answer sheets with data on background variables, etc. Their most important contribution, however, was what we called the follow-up procedure, which enabled us to study the reliability of the self-assessment.

The idea of the "follow-up" was to find out how truthful the student's answers were. This meant that the part of the test the student had done was gone through once more, but this time with the assistance of the teacher, whose task was to test the student on each item he claimed, or thought, he knew. Now it was the teacher who did the marking on the student's answer sheet according to the results of this testing procedure.

Figure 4 demonstrates the principle of the teacher's task and the possible relationships between the student's and the teacher's markings, or in other words, between self-assessment and "real" knowledge.

Figure 4.

Excerpt from answer sheet showing self-assessment and follow-up results.

Item No	S's answers Yes	?	No	
1	⊗			
2	X		○	
3	○	X		
4		X	○	
5			⊗	
Total 1	2	2	1	(Student's cross-marks)
Total 2	2	■	3	(Teacher's circles)

If the student marks the first box, thereby stating that he is absolutely certain he knows the item, and can then verify it in the follow-up, the teacher puts a circle around the cross-mark (like in Item 1). If, however, it turns out that the student does not know the answer in spite of having checked the Yes-box, the teacher puts the circle in the No-alternative (Item 2). In cases when the student marks the second option, i.e. if he is not quite sure about the correct answer, there are two possibilities for the teacher: either he puts the circle in the first box (like in Item 3), which means that the student knows the item, or in the third box (Item 4), which means he does not. Thus the second alternative is eliminated from the teacher's marking, which is logical since in reality there are only two alternatives, at least according to our criteria: either the student knows the exact answer or he does not. Finally, if the student has given a straight-forward "No" by marking the third box, the teacher always puts the circle round the student's cross-mark, which means that there is no need to test the

student on these items. The reason for this is that even if it turns out that the student has a command of such an item, it would be most probable that he has learnt it after - and maybe because of - having been confronted with the problem during the self-assessment, otherwise he would not have checked the third option.

Thanks to the different markings used by the student and the teacher on the answer sheets it was very simple to feed the results into the computer and to study the relation between self-assessment and real knowledge (as measured by the teacher), both on individual items and in more global terms.

These follow-ups were both demanding and time consuming since they involved a kind of oral testing comprising hundreds of items per student. In cases when the teacher could speak the student's native tongue or they had some other common language to communicate in, the procedure was fairly smooth. In the majority of cases, however, the follow-up had to be carried out by whatever other methods were available, e.g. by non-verbal means. In most cases the students managed to demonstrate their knowledge by using their limited Swedish proficiency, aided by gestures, paper and pencil, etc. On the whole the teachers had no serious problems in carrying out the follow-ups, but it was obvious that some types of items were more complicated or time consuming to test than others. Thus, whereas it was relatively easy to test the student's knowledge of individual words, it was rather complicated to have them explain or paraphrase the meaning of long sentences in the listening and reading comprehension subtests.

It must be emphasized that this type of follow-up is by no means meant to be part of the self-assessment process once the final version of the test has been published and made available for use, even though a certain degree of assistance and supervision on the part of the teacher will always be

desirable, if not necessary. During the developmental period, however, the follow-up procedure provides invaluable information and is therefore worth the time and effort.

5 Results and experience

Data from the try-out shows that the student's self-judgements were highly reliable as far as their yes-answers were concerned. We wish to refer to Figure 4 which can be regarded as a small scale example of the relationship between student's and teacher's scores. Although there is considerable discrepancy between them, the totals of yes-answers are the same. Similarly, in the overall results the differences between these scores were generally small. It did happen, of course, that the student could not verify the knowledge of some items for which he had checked the Yes-box. On the other hand, this loss in scores was usually compensated for by those items that the student proved to know in spite of having checked the second alternative. In most cases, however, it turned out that a mark in the second box was in reality equal to lack of knowledge. Consequently the student's and the teacher's total scores in the third column showed considerable differences. These results then seem to support our decision only to consider the yes-answers when interpreting self-assessment scores. We also wish to refer to what was said above about our cutting scores: a student is considered to have passed a difficulty level when the number of his yes-answers amounts to a minimum of 20 out of 25 items per unit. This means that even though there is normally a difference between the student's and the teacher's total scores in the yes-column, it seldom affects the final outcome, i.e. the pattern of his skills profile.

It must not be denied, however, that despite the high general agreement between the student's and teacher's total scores in the first column we had quite a few extreme cases of both over-ratings and under-ratings. Both data and teacher's

experience indicate that there is a certain correspondence between under-ratings and over-ratings on the one hand and some variables like sex, educational, national, and cultural backgrounds, proficiency level, etc. on the other. Nevertheless, we prefer to wait and see whether these preliminary results can be confirmed in a second try-out before we venture conclusions.

Correlations between self-assessment and follow-up scores were computed for all (60) units of the test battery. Due to small numbers and/or high scores in some cells these correlations must be interpreted with caution. Correlation coefficients were generally high (between .60 and .97) with the exception of the grammar sub-tests that yielded values between .45 and .65 for seven out of ten units. The teachers confirmed, on the basis of their experience from the follow-ups, that it was in the grammar sub-tests that the students made the largest number of misjudgements. This is logical since it probably applies to grammar more than to other aspects of a foreign language that what a learner thinks is correct is in actual fact incorrect.

Data from the try-out was also used to investigate the difficulty level of both individual items and entire units. It was found that we had succeeded fairly well in constructing units of increasing difficulty for each level (i.e. vertically) in spite of the fact that, not being in the possession of previous data, we had to rely on the subjective judgements of the teachers in the distribution of items to the various levels. On the other hand, we were less successful in achieving similar difficulty levels horizontally, i.e. in the various subtests belonging to the same level. The listening comprehension, for instance, was considerably easier at levels 1-6 than any other subtest at the corresponding levels. The usual pattern of skills profile achieved by our subjects, characterised by extremely high listening comprehension, did not, therefore, reflect the real state of their proficiency but was due to this deficiency in the test battery. In

developing the second version we were anxious to eliminate this shortcoming partly by re-distributing items on the basis of item analyses from the try out, partly by eliminating others and incorporating new ones.

Whether we have succeeded in our effort cannot be determined until the try-out run of the second version has been carried out. One reason why item analyses may provide insufficient information is that subjects take different sections of the test battery, which results in different samples for all levels.

As was implied above, we did not confine our evaluation to the processing of computable data. We also took stock of the experiences and subjective judgements of the teachers involved in the try-out. The stock-taking was the result of a one-week conference with the teachers during which we discussed systematically each point of a questionnaire and summarized what the teachers wished to report.

One of the most important findings was that the test was very popular with the students. They generally regarded the test as a teaching aid rather than a test, and the atmosphere while they worked on it was relaxed. They were eager to find out how far they could "climb" the scale and what sort of profile their scores would yield. They were equally anxious to find out the correct answers to all items they were uncertain about or did not know. This can be seen as a beneficial spin-off effect of the test and we think it is extremely important that the curiosity and motivation this self-assessment procedure arouses should be taken advantage of in the learning process. It was also reported that in a few cases the test did not function according to our intentions. Some students would fail to recognize the real purpose of the test, which resulted in skills profiles far above their actual proficiency. In extreme cases they refused to subject themselves to any type of follow-up procedure.

The teachers also reported that the test had met with approval by their colleagues. They particularly appreciated the opportunity of sharing with the student the responsibility for evaluation. They also thought that the diagnostic information provided by the outcome of the self-assessment could be a valuable help in their efforts to design adequate syllabuses and to individualize instruction. It was also regarded as a great advantage that, contrary to some other evaluation instruments, the content of this test could be discussed with the student, which must have a beneficial effect on learning.

6 Some problems and future plans

One of the major problems of the test is that, if taken and followed up in class, it interferes with other classroom activities. Therefore it is important that in the planning of a course sufficient time should be assigned to the use of the test. One possibility to avoid clashes between self-assessment and other classroom activities would be to allow students to do the job at home, which, of course, presupposes appropriate instructions about the test, its purpose and use. Another suggestion was that, in order to make the follow-up less time-consuming for the teacher, a complete key to test items should be made available for the students. Although this would be desirable, and also possible with several types of items, if would hardly be viable with others, expecially where the only proper key would be translation. Having a multilingual target group with more than a hundred different native languages this is practically impossible and should be limited to the vernaculars of some of the largest national groups.

Among the problems to be solved we can mention the need for a better computer program, tailored to the assessment of the difficulty levels of our test items. This is important if we are to succeed in improving the battery with regard to the difficulty levels of the various subtests.

In order to make the procedures of self-assessment and follow-up as simple as possible it is important to select through extensive and thorough experimentation the most suitable types of items. Therefore in the second version of our test we have included some new types of items which, we believe, are particularly well suited for self-assessment.

As for the use of the test it is imperative that the results should only be used for diagnostic purposes in order to assist the student in the learning process. Rewarding or penalizing students on the basis of their results would prevent them from giving honest answers and would jeopardize the beneficial effects of the procedure. We also believe that the test should be made part of a continuous diagnostic procedure, which implies that it should be taken at regular intervals during the course. In this way the test could not only provide the student and the teacher with up to date diagnostic information, but also show the student's advancement towards his goal line. However, parts of the test that have been used before, cannot be taken repeatedly without a loss of reliability. In the ideal situation there should exist two or more parallel versions, the development of which requires, of course, time and financial resources.

A try-out run of the second version of our test is now in progress and we are hoping for a sample of about 1,500 self-assessments, including 300-400 follow-ups, by June 1982. After the computation of new data as well as forthcoming conferences with teachers we shall present a final version of the test by the end of the academic year of 1982-83.

Being the result of pioneer work our final test may very well suffer from initial deficiencies. However, the primary aim of our project is not so much to produce a perfect instrument, but rather to find out whether the promising preliminary findings of our experiments with the test described in this paper can be substantiated. If so, our model will hopefully stimulate colleagues to adopt the idea of diagnostic

self-assessment tests and to construct new varieties that will best suit the special needs of their own students.

References

Oskarsson, M. 1980. Approaches to self-assessment in foreign language learning. Oxford: Pergamon.

Spolsky, B. 1976. "Language Testing: Art or Science". In Nickel, G. (ed). 1976. Proceedings of the Fourth International Congress of Applied Linguistics. Volume 3. Stuttgart: Hochschulverlag.

von Elek, T. 1977. Invandrarnas färdigheter i svenska (The Swedish proficiency of Migrants). Reports from the Language Teaching Research Center, University of Göteborg. Report No 4.

von Elek, T. (ed). 1978. Försök med alternativa modeller för undervisningen i svenska för invandrare inom AMU. (Experiments with alternative course designs in the teaching of Swedish to migrants at Labour Market Schools.) Reports from the Language Teaching Research Center, University of Göteborg. Report No 5.

von Elek, T. 1981. Självbedömning av färdigheter i svenska som andra språk. (Self-assessment of Swedish as a second language.) Work papers from the Language Teaching Research Center, University of Göteborg. Work paper No 31.

APPENDIX

Tibor von Elek
Language Teaching Research Center
University of Göteborg

April 1982

SELF-ASSESSMENT TEST

IN SWEDISH AS A SECOND LANGUAGE (SECOND VERSION)

SAMPLE ITEMS (TRANSLATION)

The test battery consists of 60 subtests or units falling into six content areas or skills (e g vocabulary, reading comprehension) on the one hand and into ten difficulty levels (1-10) on the other. Each unit can be identified by a symbol in which the letter stands for the skill tested and the digit for the level of difficulty. Thus V:1 identifies the lowest level of the vocabulary subtests and M:10 the highest of the oral units.[1] All units comprise 25 items arranged into different parts depending on the type of task they represent. The digit after each example below indicates level of difficulty.

VOCABULARY

Each of the ten units (V:1 - V:10) consists of three parts.

Part 1. Do you understand the underlined word?

Can you mend the car? (1)
He is known for his perseverance. (10)

Some other words tested in the same fashion are: *friendly (2), fetch (3), belong (5), reputation (7), exaggerate (8), bribe (9).*

Part 2. Do you know the OPPOSITE of the underlined word?

For children only. (3)
There is a shortage of teachers in this country. (9)

Some other words eliciting antonyms are: *long (1), always (2), wet (5), generous (7), optional (8), reproach (10).*

Part 3. Can you fill in the gap in this sentence?

What _____ is the dress? It's red. (3)
You shouldn't _____ to conclusions. (9)

Some other words elicited in this manner: *capital (2), blind (4), whose (5), care (7), conscience (8), /take/ measures (10).*

GRAMMAR

Each of the ten units (G:1 - G:10) consists of three parts.

Part 1. Can you supply the correct form of the word in brackets?

/BUY/ What have you _____ ? (2)
/TO WATER/ Have the flowers _____ ? (9)

Some other forms to be supplied: *child/children (1), tooth/teeth (3), careful /-ly (6), bad/worse (7).*

---

1) The letters are initials of the Swedish terms for the subtests.

Part 2. Can you fill in the gap?

Is ____ any coffee left? (1)
I look forward to ____ from you soon. (*)

Some other words elicited in this manner: *where (3), is /in a hurry/ (4), /as if he/ was/were (6), since (7), the ... the /+ the comparative/ (9),*

Part 3. Can you see what is wrong about the sentence and correct the error?

We have played football yesterday. (3)
Are you teacher? (6)
Did you think it's me? (7)

LISTENING COMPREHENSION

Each of the ten units (H:1 - H:10) consists of three parts, two of which are pictorial. All items are tape-recorded.

Part 1. Can you identify the picture we are referring to?

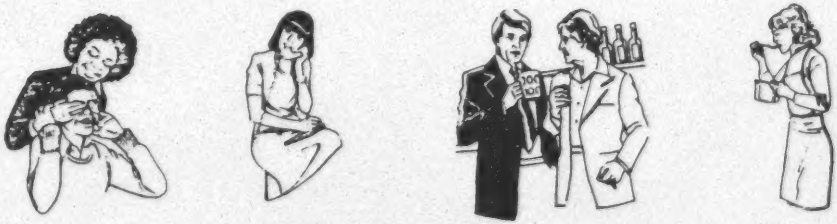
The student hears:

He is wearing glasses. (1)
My sister is talking to her boy-friend. (1)

Part 2. Can you identify the person who would be likely to utter this sentence?

The student hears:

Sorry, I must have dialled the wrong number. (4)
It's a bit too hot for my taste. (4)



(At levels 8-10 the student's task in Part 2 was to alter a picture according to oral instructions.)

Part 3. Do you understand the sentence?

Would you mind spelling your name? (3)
You can make some real bargains if you are early for the sales. (5)
According to the news agency the negotiations had to be postponed until next week. (10)

READING COMPREHENSION

Each of the ten units (L:1 - L:10) consists of at least two of the following types of items:

1. Comprehension of definitions:

Can you supply the missing word?

In _____ the goal keeper is the only person who is allowed to touch the ball. (5)
A _____ is a huge reptile. It moves awkwardly on shore but is extremely swift and dangerous in the water. (7)

2. Comprehension of non-contextual sentences:

Do you understand the sentence?

Closed on Saturdays. (1)
There are two exceptions to this rule. (4)
Due to our well-organised pre- and post-natal care the infant mortality rate is extremely low. (8)

3. Comprehension of coherent texts.

The number, length and complexity of the texts varied considerably at different levels. Each text was followed by a number of questions concerning words, expressions or context.

ORAL PRODUCTION

Each of the ten units (M:1 - M:10) consists of at least three of the following tape-recorded parts.

1. Repetition of sentences.

Can you repeat the sentence?

See you tomorrow! (1)
This is my daughter Maria. She is eleven. (2)
If someone had told me ten years ago that I would eventually settle down in Sweden I would have taken it as a joke. (9)

2. Answering questions.

Can you give an intelligible answer to the question?

How long have you been in Sweden? (2)
Can you give some reasons why people emigrate? (6)
What kind of reform would you like to introduce in your country and why? (9)

3. Picture-cued tasks.

Look at Picture A. What is the weather like here? (1)
Look at sceneries A and B. Which of these places would you choose for your holidays and why? (5)
Choose one of these pictures. Could you describe it so closely that a person who does not see the picture could get a good idea of what it represents? (10)

A

B

4. Coping with situations.

A Swede tells you that he is about to leave for your native country as a tourist. Could you advise him what he should see and do there? (4)
You are applying for a job. Could you tell the agent about your qualifications and previous experience? (6)
Someone claims that immigrants cause Sweden considerable economic losses. Could you defeat this view using objective arguments? (9)

5. Definitions, comparisons

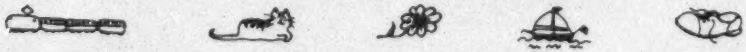
Could you explain the difference between a restaurant and a café? (4)
Could you explain the similarities and differences between a tiger and a lion? (6)
Could you explain to someone what superstition means? (10)

WRITTEN PRODUCTION

Each of the ten units (S:1 - S:10) comprises at least three of the following parts.

1. Spelling

Can you spell the following words? (2)



Can you discover the spelling mistakes in this sentence and correct them?
I am particularly interested in litterature and classical music. (10)

2. Morphology and syntax.

Can you produce correct sentences of these chains of words?

daughter / ill / cannot / go / school. (4)
grateful / information / when / product / deliver. (5)

3. Completion of sentences.

Could you finish the sentence?

Thank you for your invitation. I very much regret ... (6)
With reference to your inquiry ... (10)

4. Coping with situations.

Can you write a short autobiography? (6)
You have read an article in your native language. You think some of your Swedish friends should read it, too. Could you translate the article or write a resumé containing its essential points. (9)

5. Definitions, comparisons.

Could you explain the difference between an emigrant and a refugee in a few lines? (5)
Could you define the word hostage in a few lines? (9)

REPORTS FROM THE LANGUAGE TEACHING RESEARCH CENTER

1 Nelhans, B. Lagstadgad undervisning i svenska för invandrare: En sociologisk utredning om lagen om rätt till ledighet och lön vid deltagande i svensk-undervisning för invandrare. (1975)

2 Blåberg, C. Engelska på vuxengymnasiet: En jämförelse mellan en- och tvåkvällskurser. (1977)

3 Josefson, I. Lagstadgad undervisning i svenska för invandrare: Mål, metoder, material. (1977)

4 von Elek, T. Invandrarnas färdigheter i svenska: Rapport från Projektet AMU-Test. (1977)

5 von Elek, T. (red.) Försök med alternativa modeller för undervisningen i Svenska för invandrare inom AMU: Rapport från Projektet AMU-Sfi. (1978)

6 Oskarsson, M. Språkcirkelledares behov av fortbildning: Rapport om en undersökning inom Studieförbundet Vuxenskolan. (1979)

7 Oskarsson, M. Approaches to Self-Assessment in Foreign Language Learning. (1979)

WORK PAPERS FROM THE LANGUAGE TEACHING RESEARCH CENTER

1 Oskarsson, M. The Acquisition of Foreign Language Grammar by Adults: A Summary Report on Three Field Experiments. Paper read at the Third International Congress of Applied Linguistics, Copenhagen 1972 (16 pp., 1974)

2 von Elek, T. Språktest: En exempelsamling (14 s., 1974)

3 Oskarsson, M. Mätning av språkkunskaper: En exempelsamling med arbetsuppgifter (24 s., 1974)

4 Oskarsson, M. Monolingual and Bilingual Vocabulary Learning: Theoretical Background and Experimental Findings. Paper read at the IATEFL (International Association of Teachers of English as a Foreign Language) Conference, Budapest 1974 (39 pp., 1975)

5 Oskarsson, M. Ett statistiskt försvar för flervalstest (6 s., 1975)

6 von Elek, T. Ett experiment med individualiserad grammatikundervisning (10 s., 1975) Finns även på engelska.

7 von Elek, T. Experiments in Teaching Foreign Language Grammar by Different Methods: A Summing-Up of the Research of the GUME/Adults Project. Paper read at the Fourth International Congress of Applied Linguistics, Stuttgart 1975 (11 pp., 1976)

8 Oskarsson, M. The Relationship between Foreign Language Proficiency and Various Psychological Variables. Paper read at the Fourth International Congress of Applied Linguistics, Stuttgart 1975 (15 pp., 1976)

9 Oskarsson, M. Svenska för invandrare och projektet Språk-Vux. Föredrag inför Association Norvégienne de Linguistique Appliquée, Bergen 1976 (6 pp., 1976)

10 Oskarsson, M. Guided Self-Assessment in Adult Language Learning: Some notes on the Situation in Sweden and a Research Proposal. Council of Europe, Strasbourg (5pp., 1976)

11 Oskarsson, M. Moderna språk i vuxen- och invandrarundervisningen: Rapport om Projektet Språk-Vux (11 s., 1976)

12 Oskarsson, M. Hur lär vi oss språk? Några allmänna synpunkter på språkinlärning (9 s., 1977)

13 Oskarsson, M. Språkpedagogisk forskning: En beskrivning av några projekt (17 s., 1977)

14 von Elek, T. Språkfärdigheten i svenska hos vuxna invandrare: Undersökningar inom Språkpedagogiskt Centrum (12 s., 1978)

15 Oskarsson, M. Objective Formalized Assessment versus Subjective Self-Assessment: Two Approaches to the Evaluation of Oral Proficiency. Paper read at the British Council/Goethe Institut Colloquium on Evaluating Oral Production, Paris 1978 (12 pp., 1978)

16 Oskarsson, M. et al., Försök med behovsanpassad undervisning och självbedömning av färdigheter i moderna språk (49 s., 1978) (Utgår. Ersätts av Nr 28)

17 Oskarsson, M. Experiments on Self-Assessment in Foreign Language Learning. Paper read at the Fifth International Congress of Applied Linguistics, Montreal 1978 (9 pp., 1978)

18 Oskarsson, M. En enkätundersökning om försöksverksamhet kring vuxnas inlärning av främmande språk (25 s., 1978)

19 Oskarsson, M. Grammatiska termer på svenska och engelska: En lista med exempel och definitioner (7 s., 1979)

20 Oskarsson, M. Om ord, uttal och grammatik i vuxnas språkinlärning (7 s., 1979)

21 Oskarsson, M. Språkcirkelledares behov av fortbildning: Sammanfattning av en undersökningsrapport (9 s., 1979)

22 Oskarsson, M. Grammatical Terms in English and Swedish: A List with Examples and Definitions (8 pp., 1979)

23 Andersson, E. Boken och läsaren: Något om lättlästa böcker och hur böcker blir lättlästa. (70 s., 1979)

24 Oskarsson, M. Subjective and Objective Assessment of Foreign Language Performance: Discussion and Experimental Findings. Paper read at The Regional Seminar on Evaluation and Measurement of Language Competence and Performance, Singapore 1980 (24 pp., 1980)

25 Jönsson, E. Försök med muntligt inriktade språkkurser i kommunal vuxenutbildning (131 s., 1980)

26 Malmström, L. Studiecirkeldeltagares behov av språkkunskaper (188 s., 1980)

27 Carlsson, N. En grupp vuxenelevers syn på språkliga behov och nyss genomgången språkutbildning (44 s., 1980)

28 Oskarsson, M. Behovsanpassad undervisning och självbedömning av färdigheter i moderna språk: Några försök (79 s., 1980)

29 Malmström, L. Försök med kommunikativt inriktade övningar i engelska språkcirklar (160 s., 1981)

30 von Elek, T. Självbedömning av färdigheter i svenska som andra språk: Delrapport från ett utvecklingsprojekt (149 s., 1981)

31 von Elek, T. Test of Swedish as a Second Language: An Experiment in Self-Assessment (26 pp., 1982)